

**[DESCRIPTION]**

**[Invention Title]**

**APPARATUS AND METHOD FOR EXTRACTING REPRESENTATIVE  
STILL IMAGES FROM MPEG VIDEO**

5      **[Technical Field]**

The present invention relates generally to an apparatus and method for extracting representative still images from Moving Picture Experts Group video and, more particularly, to an apparatus and method for extracting 10 representative still images from Moving Picture Experts Group video, which extracts representative still images from Moving Picture Experts Group 1 or Moving Picture Experts Group 2 video, and provides the extracted representative still images to a user at high speed.

15      **[Background Art]**

With recently developed digital technology, multimedia data, such as high-quality video or music, can be generated more easily and quickly than before. Generally, such multimedia data is characterized in that 20 required storage capacity is considerably large and playing time is considerably long. Accordingly, in order to efficiently store, search and read such multimedia data, various technologies have been required, and related research and efforts have been conducted. As a result, the

size of such multimedia data can be considerably reduced through international compression standards, such as Moving Picture Experts Group (MPEG) 1 and MPEG 2 standards, and research into MPEG 7, which allows multimedia data to be 5 efficiently read and searched, is being conducted.

In particular, a technology, which allows video having long playing time to be read at high speed, is referred to as a "video abstract." A video abstract formed of still images is referred to as a "video summary," and a 10 video abstract including video and related audio information is referred to as "video skimming."

Since the video summary uses only still images, the video summary is characterized in that it can be generated faster than the video skimming. On the other hand, the 15 video skimming is characterized in that it can provide more natural screens to a user using audio and textural information.

The video summary is a set of representative still images that represents the contents of video desirably, and 20 the methods thereof are classified according to how to select the representative still images.

A method of extracting representative still images at regular periods is disadvantageous in that some of the representative still images may be missed because the 25 representative still images are not distributed at regular intervals.

A method of extracting one still image for each shot

of video is disadvantageous in that the number of representative still images and temporal distribution are determined by the number of shots and the temporal distribution. That is, an excessively large number of still 5 images or a very small number of still images may be selected according to the number of shots.

Such conventional methods of extracting various feature values from video and nonlinearly extracting representative still images from a feature space are 10 characterized in that calculating time is long or irregular according to the variation in the contents of the video.

The conventional methods as described above are problematic in that processing time is too long to provide a video summary to a user at high speed, or in that it is 15 difficult to predict the processing time.

#### **[Disclosure]**

#### **[Technical Problem]**

Accordingly, an object of the present invention is to provide an apparatus and method for extracting 20 representative still images from MPEG video, which can provide a desired number of representative still images at high speed and predictable speed when the representative still images for generating a video summary are extracted from the MPEG video.

#### **25 [Technical Solution]**

In order to accomplish an object, the present invention provides an apparatus for extracting representative still images from MPEG video, including a video curve generation unit for calculating distances between adjacent frames of all intra frames of input video and generating a video curve that is a cumulative curve of the distances; a video curve division unit for dividing the video curve into a certain number of segments; a still image selection unit for selecting video images corresponding to certain points of the divided video curve as representative still images; and a video output unit for outputting the still images selected by the still image generation unit.

Furthermore, the video curve generation unit may include an intra frame selection unit for selecting an intra frame from the input video; at least one Y picture selection unit for selecting only Direct Current (DC) coefficients from Discrete Cosine Transform (DCT) coefficients of a Y picture on the selected intra frame; at least one cumulative DC histogram generation unit for extracting a cumulative histogram of the DC coefficients; at least one frame distance generation unit for calculating a maximum distance between cumulative histograms of adjacent intra frames and determining the maximum distance to be a distance between two adjacent frames; and a cumulative frame distance histogram generation unit for acquiring the video curve, that is, a cumulative curve,

from the distance between the adjacent frames of the selected intra frames when the distance between the adjacent frames is calculated through the Y picture selection unit, the cumulative DC histogram generation unit  
5 and the frame distance generation unit.

In order to accomplish another object, the present invention provides a method of extracting representative still images from MPEG video, including the steps of generating a video curve, that is, a cumulative curve of  
10 distances between adjacent frames of all frames of input video, by calculating the distances between the frames; dividing the video curve into a certain number of segments; selecting video images corresponding to certain points of the divided video curve as representative still images; and  
15 outputting all or some of the selected still images.

Furthermore, the step of generating the video curve may include the steps of selecting an intra frame of the input video; selecting only DC coefficients from DCT coefficients of a Y picture on the selected intra frame;  
20 extracting a cumulative histogram of the DC coefficients; calculating a maximum distance between cumulative histograms of adjacent intra frames and determining the maximum distance to be a distance between two neighboring frames; and acquiring the video curve, that is, a cumulative curve of distances between neighboring frames of all selected intra frames, by calculating the distances  
25 between the adjacent frames.

**[Description of Drawings]**

FIG. 1 is a view showing the construction of an apparatus for extracting representative still images from MPEG video according to an embodiment of the present invention;

FIG. 2 is a detailed view showing the construction of the video curve generation unit of FIG. 1;

FIG. 3 is a graph showing an example of the division of a video curve according to an embodiment of the present invention;

FIG. 4 is graphs showing the approximation line and tangent point of the video curve according to an embodiment of the present invention;

FIG. 5 is a view showing an example of still images output according to an embodiment of the present invention; and

FIG. 6 is a flowchart showing a method of extracting representative still images from MPEG video according to an embodiment of the present invention.

**20 [Best Mode]**

The present invention is described in more detail through embodiments below. The embodiments are described only for illustrative purposes, and the scope of the present invention is not limited by the embodiments.

FIG. 1 is a view showing the construction of an

apparatus for extracting representative still images from MPEG video according to an embodiment of the present invention.

The embodiment of the present invention is described  
5 with reference to FIG. 1 below.

The apparatus for extracting representative still images according to the present invention includes a video curve generation unit 100 for acquiring a video curve from MPEG 1 or MPEG 2 video, a video curve division unit 200 for dividing the video curve into n segments (n is a natural number), a still image generation unit 300 for selecting a video scene, which corresponds to an nth-order approximation tangent point of the divided video curve, as an nth still image, and a video output unit for exhibiting  
10 selected n still images.  
15

The video curve generation unit (video curve extraction apparatus) 100 extracts a video curve from MPEG 1 or MPEG 2 video.

After the video curve has been extracted, the video  
20 curve division unit 200 receives a desired number of representative still images from a user through a user requirement input unit 210 and divides the video curve by the desired number.

The still image generation unit 300 selects a video  
25 scene, which corresponds to an nth-order approximation tangent point of the divided video curve, as a representative still image. The number of representative

still images may be a certain number equal to or less than n.

The video output unit 400 exhibits the still images, which are selected by the still image generation unit, to 5 the user.

FIG. 2 is a detailed view showing the video curve generation unit 100 shown in FIG. 1.

Referring to FIG. 2, the video curve generation unit (video curve extraction apparatus) 100 includes an intra frame selection unit 110, a Y picture selection unit 120, a cumulative DC histogram generation unit 130, a frame distance generation unit 140, and a cumulative frame distance histogram generation unit 150.

The intra frame selection unit 110 selects only the 15 intra frames of MPEG video from input MPEG video.

The Y picture selection unit 120 selects only DC coefficients from the DCT coefficients of a Y picture on the selected intra frames.

The cumulative DC histogram generation unit 130 20 extracts a cumulative histogram of the selected DC coefficients. In the cumulative histogram generation unit 130, the DC histogram is the frequency distribution of the DC values of pixels in the video. If it is assumed that an nth frequency value (i.e., a DC histogram value) is H\_DC(n), a cumulative histogram cH\_DC(n) is determined by 25 values H\_DC(n-1) and H\_DC(n).

The frame distance generation unit 140 calculates the

maximum distance between the cumulative histogram and an adjacent cumulative histogram, and determines the calculated distance to be a distance between adjacent frames. In the frame distance generation unit 140, a frame 5 distance refers to the largest value acquired by calculating the differences between the Y-axial values of two histograms according to frequencies if it is assumed that the cumulative DC histograms of neighboring still images are  $cH\_DC(n-1)$  and  $cH\_DC(n)$ , respectively.

When the distances between the selected intra frames are calculated through the units 120 to 140, the cumulative frame distance histogram generation unit 150 acquires a cumulative curve from the distances. The cumulative curve is referred to as a "video curve." In the cumulative frame 15 distance histogram generation unit 150, the cumulative frame distance histogram is acquired by cumulating the calculated frame distances using the same method used by the cumulative DC histogram generation unit 130.

FIG. 3 is a graph showing an example of the division 20 of a video curve according to an embodiment of the present invention.

Referring to FIG. 3, the X-axis of the video curve corresponds to the time axis of video, and the Y-axis of the video curve corresponds to the distance between the 25 cumulative frames.

Slopes at points of the video curve are proportional to the amount of variation in contents between frames. That

is, a high-sloped section indicates the case where visual variation is very large in the video. Furthermore, a low-sloped section indicates the case where visual variation is very small in the video.

5 In the video curve division unit 200, the video curve acquired as described above is divided into n segments (n is a natural number). Furthermore, desired n representative still images are received from a user and the video curve may be divided into n segments.

10 A line formed by connecting two end points of the video curve within each divided segment is referred to as an "approximation line." Furthermore, it is assumed that "the division of the video curve into n segments" is "the video curve is approximated with n approximation lines."

15 Since a division method is closely related to the acquisition of an approximation tangent point, a description of the division method is made in conjunction with the approximation tangent point below.

FIG. 4 is graphs showing the approximation line and 20 tangent point of the video curve according to an embodiment of the present invention.

In the following description, a line formed by connecting a start point and an end point is referred to as a first-order approximation line. In FIG. 4, the first-order approximation line is a line formed by connecting the 25 start point of a curve placed at the lowermost left corner with the end point of the curve placed at the uppermost

right corner. A point on the curve farthest from the first-order approximation line is a first-order approximation tangent point. A video screen corresponding to the first-order approximation tangent point is referred to as a  
5 "first representative still image." Second-order approximation lines are lines placed on both sides of the first approximation tangent point, so that the second-order approximation lines are two. A second-order approximation tangent point is a point on the curve farthest from the two  
10 second-order approximation lines. A video screen corresponding to the second-order approximation tangent point is referred to as a "second representative still image." FIG. 4 shows an example of the second-order approximation lines, the second-order approximation tangent  
15 point d<sub>2</sub>, a third-order approximation line and a third-order approximation tangent point d<sub>3</sub>.

The video curve is divided into n segments using the above-described method, and an nth-order approximation line, an nth-order approximation tangent point and an nth  
20 representative still image are acquired.

The maximum values of the differences between approximation lines and the Y-axial values of the video curve are obtained to match a desired number of representative still images using the method as described  
25 above, thus finally finding n representative still images.

The video curve according to the embodiment of the present invention is a rising curve the slope of which is

greater than '0'. In the case of the rising curve, the distance between a video curve and an approximation line can be simply calculated using the difference between Y-axial values.

5 Furthermore, the present invention is characterized in that a path, which must be scanned to acquire the distance between an approximation line and a video curve, is not related to the slope variation of the curve but is proportional only to the playing time of the video, that is, the number of intra frames. That is, the scanned path is regularly maintained even though scanning is performed repeatedly, so that entire scanned paths, which are scanned until the video curve is approximated with n lines, are not related to the slope variation of the video curve but are 10 linearly increased in proportion to the number n and the length of an X axis.

15

The video output unit 300 outputs the still images acquired as described above.

FIG. 5 is a view showing an example of output still 20 images according to the embodiment of the present invention.

FIG. 6 is a flowchart showing a method of extracting video still images according to the embodiment of the present invention.

25 Referring to FIG. 6, the method of extracting video still images from video is performed in such a way as to select an intra frame from MPEG video at step 510. Only DC

coefficients are selected from the DCT coefficients of a Y picture on the selected intra frame at step 520. A cumulative histogram of the DC coefficients is extracted at step 530. The maximum distance between two adjacent 5 cumulative histograms is calculated and determined to be the distance between two intra frames at step 540. A video curve (cumulative curve) is acquired by calculating the distance between the frames at step 550. The video curve is divided into n segments (n is a natural number) at step 10 560. A video scene corresponding to the nth-order approximation tangent point of the divided video curve (n is a natural number) is selected as an nth still image at step 570. The n still images are shown at step 580. All or some of still images can be shown according to necessity.

15 In the above description, steps 510 to 550 are the steps of extracting a video curve from video.

#### **[Industrial Applicability]**

As described above, the present invention can generate a video summary from video, more particularly, 20 from MPEG 1 or MPEG 2 video, and provide the generated video summary to the user at high speed.

Furthermore, the present invention can provide a desired number of representative still images to the user.

Furthermore, since processing time is not related to 25 the variation in the contents of video but is proportional to the number of still images desired by the user, it is

possible to predict waiting time for the video summary.